

OIPE

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/902,693

DATE: 12/07/2001

TIME: 16:34:31

Input Set : N:\Crf3\RULE60\09902693.txt

Output Set : N:\CRF3\12072001\I902693.raw

3 <110> APPLICANT: MIHARA, Yasushiro  
4 TAKEUCHI, Sonoko  
5 JOJIMA, Yasuko  
6 TONOUCHI, Naoto  
7 FUDOU, Ryosuke  
8 YOKOZEKI, Kenzo  
10 <120> TITLE OF INVENTION: NOVEL MICROORGANISM AND METHOD FOR PRODUCING XYLITOL OR  
11 D-XYLULOSE  
13 <130> FILE REFERENCE: 0010-1015-0  
15 <140> CURRENT APPLICATION NUMBER: 09/902,693  
16 <141> CURRENT FILING DATE: 2001-07-12  
18 <150> PRIOR APPLICATION NUMBER: 09/347,001  
19 <151> PRIOR FILING DATE: 1999-07-02  
21 <150> PRIOR APPLICATION NUMBER: JP 10-193472  
22 <151> PRIOR FILING DATE: 1998-07-08  
24 <150> PRIOR APPLICATION NUMBER: JP 10-310398  
25 <151> PRIOR FILING DATE: 1998-10-30  
27 <150> PRIOR APPLICATION NUMBER: JP 11-12244  
28 <151> PRIOR FILING DATE: 1999-01-20  
30 <160> NUMBER OF SEQ ID NOS: 5  
32 <170> SOFTWARE: PatentIn Ver. 2.1  
34 <210> SEQ ID NO: 1  
35 <211> LENGTH: 1438  
36 <212> TYPE: DNA  
37 <213> ORGANISM: Unknown Organism  
39 <220> FEATURE:  
40 <223> OTHER INFORMATION: Description of Unknown Organism: strain P528  
42 <220> FEATURE:  
43 <223> OTHER INFORMATION: N at position 1365 is A, T, G, or C  
45 <400> SEQUENCE: 1  
46 tgatcctggc tcagagcgaa cgetggcggc atgcttaaca catgcaagtc gcacggacct 60  
47 ttccgggtga gtggcggaag ggtgagtaac gcgtaggat ctatccacgg gtgggggata 120  
48 aactgggaa actggtgcta ataccgcatg atacctgagg gtcaaaggcg cgagtcgcct 180  
49 gtggaggagc ctgcgttcga ttagcttgtt ggtggggtaa aggcctacca aggcgatgat 240  
50 cgatagctgg tctgagagga tgatcagcca cactgggact gagacacggc ccagactcct 300  
51 acgggaggca gcagtgggga atattggaca atgggcgcaa gctgatcca gcaatgccgc 360  
52 gtgtgtgaag aaggtcttcg gattgtaaag cactttcgac ggggacgatg atgacggtac 420  
53 ccctagaaga agcccccggc aacttcgtgc cagcagccgc ggtaatacga agggggctag 480  
54 cgttgctcgg aatgactggg cgtaaagggc gtgtaggcgg ttgttacagt cagatgtgaa 540  
55 attccagggc ttaaccttgg ggtgcattt gatacgtagc gactagagtg tgagagaggg 600  
56 ttgtggaatt cccagtgtag aggtgaaatt cgtagatatt ggggaagaaca ccggtggcga 660  
57 aggcggcaac ctggtcctatg actgacgctg aggcgcgaaa gcgtggggag caaacaggat 720  
58 taataccct ggtagtccac gctgtaaagc atgtgtgctg gatgttgggt aacttagtta 780  
59 ctacgtgtcg aagctaacgc gctaagcaca ccgcctggga agtacggccg caaggttgaa 840  
60 actcaaagga attgaacggg gcccgacaaa gcggtggagc atgtggttta attcgaagca 900  
61 acgcgcagaa ccttaaccagg gcttgacatg gggaggctgt actcagagat gggtatattcc 960  
62 cgcaagggac ctctgcaca ggtgctgcat ggctgtcgtc agctcgtgtc gtgagatgtt 1020

ENTERED

## RAW SEQUENCE LISTING

DATE: 12/07/2001

PATENT APPLICATION: US/09/902,693

TIME: 16:34:31

Input Set : N:\Crf3\RULE60\09902693.txt

Output Set: N:\CRF3\12072001\I902693.raw

```

63 ggggttaagtc ccgcaacgag cgcaaccctc gccttttagtt gccagcacgt ttgggtgggc 1080
64 actctagagg aactgccggg gacaagccgg aggaagggtg ggatgacgtc aagtcctcat 1140
65 ggccttatatg tcctgggcta cacaagtgtt acaatggcgg tgacagtggg aagctagatg 1200
66 gtgacatcat gccgatctca aaaagccgtc tcagttcgga ttgtactctg caactcgagt 1260
67 acatgaaggt ggaatcgcta gtaatcgagg atcagcatgc cgcggtgaat acgttccccg 1320
W--> 68 gccttgtaga caccgccccg cacaccatgg gagttgggtt gaccngaagc cggtagagca 1380
69 accgcaagga cgcagccgac caccggtcggg tcagcgactg gggtagagtc gtaacaag 1438
72 <210> SEQ ID NO: 2
73 <211> LENGTH: 1436
74 <212> TYPE: DNA
75 <213> ORGANISM: Unknown Organism
77 <220> FEATURE:
78 <223> OTHER INFORMATION: Description of Unknown Organism:strain S877
80 <400> SEQUENCE: 2
81 tgatcctggc tcagagcgaa cgctggcggc atgcttaaca catgcaagtc gcacgaacct 60
82 ttccgggtta gtggcgagcg ggtgagtaac gcgtaggaac ctatccagag gtgggggata 120
83 acaccgggaa actggtgcta ataccgcatg atacctgagg gttaaaggct tttgttgctt 180
84 ttggaggggc ctgcgtttga ttagctagtt ggttgggtaa aggctgacca aggcgatgat 240
85 caatagctgg tttgagagga tgatcagcca cactgggact gagacacggc ccagactcct 300
86 acgggaggca gcagtgggga atattggaca atgggggcaa ccctgatcca gcaatgccgc 360
87 gtgtgtgaag aaggctcttcg gattgtaaag cactttcact aggggaagatg atgacgggtac 420
88 ctagagaaga agccccggtt aacttcgtgc cagcagccgc ggtaatacga agggggctag 480
89 cgttgctcgg aatgactggg cgtaaagggc gcgtaggcgg ttatacagt cagatgtgaa 540
90 atccccgggc ttaacctggg aactgcattt gatacgtata gactagagtc cgagagagga 600
91 ttgcggaatt cccagtgtag aggtgaaatt cgtagatatt gggaagaaca ccagttgcga 660
92 aggcggcaat ctggctcgga actgacgtg aggcgcgaaa gcgtggggag cgaacaggat 720
93 tagataacct ggtagtccac gctgtaaacg atgtgtgctg gatgttggga aacttagttt 780
94 ttcagtgtcg aagctaactg gttaagcaca ccgcctgggg agtacgaccg caagggttga 840
95 actcaaagaa attgacgggg gccgcacaaa gcggtggagc atgtggttta attcgaagca 900
96 acgcgcagaa ccttaccagg tcttgtatgg ggaggacgtg ctacagatg agtatttctt 960
97 cggacctccc gcacagggtg tgcatggctg tcgtcagctc gtgtcgtgag atgttgggtt 1020
98 aagtcgccga acgagcgcaa cccctgtctt tagttgccat cagtttggg tgggcactct 1080
99 agagagactg ccggtgacaa gccggaggaa ggtggggatg acgtcaagtc ctcatggccc 1140
100 ttatgacctg ggctacacac gtgctacaat ggcggtgaca atgggaagct acatggtgac 1200
101 atgatgccga tctcaaaaaa ccgtctcagt tcggattgca ctctgcaact cgagtgcagt 1260
102 aaggtggaat cgctagtaat cgtggatcag catgccacgg tgaatacgtt cccgggcctt 1320
103 qtacacaccg cccgtcacac catgggagtt ggtttgacct taagccggtg agcgaaccgc 1380
104 aagggcgcag cgacccacgg tcgggtcagc gactggggtg aagtcgtaac aaggta 1436
107 <210> SEQ ID NO: 3
108 <211> LENGTH: 691
109 <212> TYPE: DNA
110 <213> ORGANISM: Unknown Organism
112 <220> FEATURE:
113 <223> OTHER INFORMATION: Description of Unknown Organism:strain S1009
115 <400> SEQUENCE: 3
116 tgatcctggc tcagagcgaa cgctggcggc atgcttaaca catgcaagtc gcacgaacct 60
117 ttccgggtta gtggcgagcg ggtgagtaac gcgtaggaac ctatccagag gtgggggata 120
118 acaccgggaa actggtgcta ataccgcatg atacctgagg gttaaaggct tttgttgctt 180
119 ttggaggggc ctgcgtttga ttagctagtt ggttgggtaa aggctgacca aggcgatgat 240

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/902,693

DATE: 12/07/2001

TIME: 16:34:31

Input Set : N:\Crf3\RULE60\09902693.txt

Output Set : N:\CRF3\12072001\I902693.raw

```

120 caatagctgg tttgagagga tgatcagcca cactgggact gagacacggc ccagactcct 300
121 acgggaggca gcagtgggga atattggaca atgggggcaa ccttgatcca gcaatgccgc 360
122 gtgtgtgaag aagggtcttcg gattgtaaag cactttcact agggaagatg atgacggtac 420
123 ctagagaaga agccccggct aacttcgtgc cagcagccgc ggtaatacga agggggctag 480
124 cgttgctcgg aatgactggg cgtaaagggc gcgtaggcgg tttatacagt cagatgtgaa 540
125 atccccgggc ttaacctggg aactgcattt gatacgtata gactagagtc cgagagagga 600
126 ttgcggaatt cccagtgtag aggtgaaatt cgtagatatt gggaagaaca ccagttgcga 660
127 aggcggcaat ctggctcgga actgacgctg a 691

```

130 &lt;210&gt; SEQ ID NO: 4

131 &lt;211&gt; LENGTH: 691

132 &lt;212&gt; TYPE: DNA

133 &lt;213&gt; ORGANISM: Unknown Organism

135 &lt;220&gt; FEATURE:

136 &lt;223&gt; OTHER INFORMATION: Description of Unknown Organism:strain S1019

138 &lt;400&gt; SEQUENCE: 4

```

139 tgatcctggc tcagagcgaa cgctggcggc atgcttaaca catgcaagtc gcacgaacct 60
140 ttccgggtta gtggcggacg ggtgagtaac gcgtaggaa cctatccagag gtgggggata 120
141 acaccgggaa actggtgcta ataccgcatg atacctgagg gttaaaggct tttgttgctt 180
142 ttggaggggc ctgcgtttga ttagctagtt ggttgggtaa aggctgacca aggcgatgat 240
143 caatagctgg tttgagagga tgatcagcca cactgggact gagacacggc ccagactcct 300
144 acgggaggca gcagtgggga atattggaca atgggggcaa ccttgatcca gcaatgccgc 360
145 gtgtgtgaag aagggtcttcg gattgtaaag cactttcact agggaagatg atgacggtac 420
146 ctagagaaga agccccggct aacttcgtgc cagcagccgc ggtaatacga agggggctag 480
147 cgttgctcgg aatgactggg cgtaaagggc gcgtaggcgg tttatacagt cagatgtgaa 540
148 atccccgggc ttaacctggg aactgcattt gatacgtata gactagagtc cgagagagga 600
149 ttgcggaatt cccagtgtag aggtgaaatt cgtagatatt gggaagaaca ccagttgcga 660
150 aggcggcaat ctggctcgga actgacgctg a 691

```

153 &lt;210&gt; SEQ ID NO: 5

154 &lt;211&gt; LENGTH: 691

155 &lt;212&gt; TYPE: DNA

156 &lt;213&gt; ORGANISM: Unknown Organism

158 &lt;220&gt; FEATURE:

159 &lt;223&gt; OTHER INFORMATION: Description of Unknown Organism:strain S1023

161 &lt;400&gt; SEQUENCE: 5

```

162 tgatcctggc tcagagcgaa cgctggcggc atgcttaaca catgcaagtc gcacgaacct 60
163 ttccgggtta gtggcggacg ggtgagtaac gcgtaggaa cctatccagag gtgggggata 120
164 acactgggaa actggtgcta ataccgcatg atacctgagg gtcaaaggct tttgttgctt 180
165 taggaggggc ctgcgtttga ttagctagtt ggttgggtaa aggctgacca aggcgatgat 240
166 caatagctgg tttgagagga tgatcagcca cactgggact gagacacggc ccagactcct 300
167 acgggaggca gcagtgggga atattggaca atgggggcaa ccttgatcca gcaatgccgc 360
168 gtgtgtgaag aagggtcttcg gattgtaaag cactttcact agggaagatg atgacggtac 420
169 ctagagaaga agccccggct aacttcgtgc cagcagccgc ggtaatacga agggggctag 480
170 cgttgctcgg aatgactggg cgtaaagggc gcgtaggcgg tttatacagt cagatgtgaa 540
171 atccccgggc ttaacctggg aactgcattt gatacgtata gactagagtc cgagagagga 600
172 ttgcggaatt cccagtgtag aggtgaaatt cgtagatatt gggaagaaca ccagttgcga 660
173 aggcggcaat ctggctcgga actgacgctg a 691

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/902,693

DATE: 12/07/2001

TIME: 16:34:32

Input Set : N:\Crf3\RULE60\09902693.txt

Output Set: N:\CRF3\12072001\I902693.raw

L:68 M:258 W: Mandatory Feature missing, <221> not found for SEQ ID#:1

L:68 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:1

L:68 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1